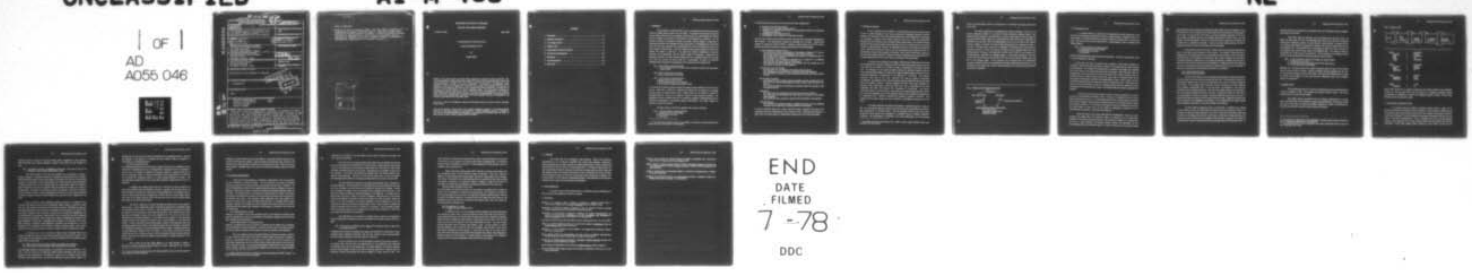
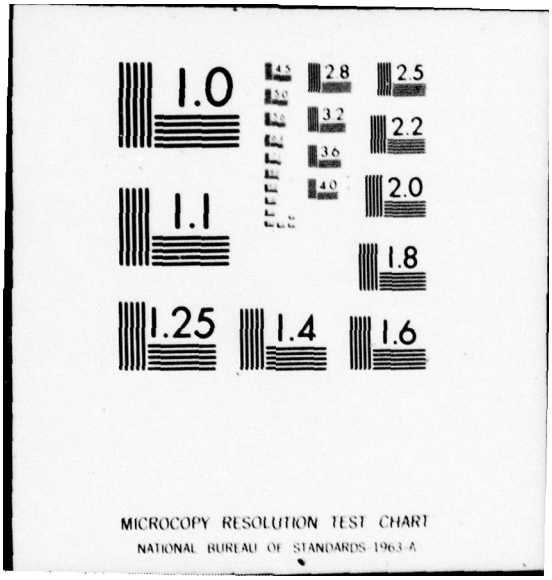
AD-A055 046 MASSACHUSETTS INST OF TECH CAMBRIDGE ARTIFICIAL INTE--ETC F/G 9/2
A PROGRESS REPORT ON THE DISCOURSE AND REFERENCE COMPONENTS OF --ETC(U)
APR 78 C SIDNER N00014-75-C-0643
UNCLASSIFIED AI-M-468 NL

| OF |
AD
A055 046



END
DATE
FILMED
7-78
DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

FOR FURTHER TRAN

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AIM-468	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): A Progress Report on the Discourse and Reference Components of PAL	5. TYPE OF REPORT & PERIOD COVERED: memo
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Candace Sidner	8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0643
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, Massachusetts 02139	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, 1400 Wilson Blvd, Arlington, Virginia 22209	12. REPORT DATE: Apr 78
	13. NUMBER OF PAGES: 190
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Office of Naval Research, Information Systems, Arlington, Virginia 22217	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

This document has been approved for public release and sale; its distribution is unlimited.

DDC RECEIVED JUN 14 1978 F

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

UNCLASSIFIED

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Reference disambiguation
Discourse interpretation,
Discourse purposes
Natural language
Focus
Frames

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper reports on research being conducted on a computer assistant, called PAL. PAL is being designed to arrange various kinds of events with concern for the who, what , when, where and why of that event. The goal for PAL is to permit a speaker to interact with it in English and to use extended discourse to state the speaker's requirments. The portion of the language system discussed in this report disambiguates references from discourse and interprets the purpose of sentences of the discourse. PAL uses the focus of discourse to direct its attention to a portion of the discourse and to the

over

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

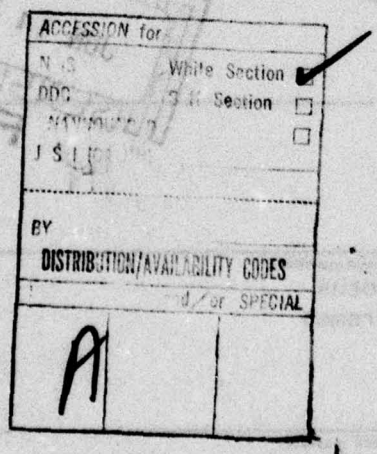
UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

407 483

BLOCK 20 CONTINUED

database to which the discourse refers.. The focus makes it possible to disambiguate references with minimal search. Focus and a frames representation of the discourse make it possible to interpret discourse purposes. The focus and representation of the discourse are explained, and the computational components of PAL which implement reference disambiguation and discourse interpretation are presented in detail.



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

ARTIFICIAL INTELLIGENCE LABORATORY

AI Memo No. 468

April 1978

A Progress Report on the Discourse and Reference Components of PAL

by

Candace Sidner

Abstract: This paper reports on research being conducted on a computer assistant, called PAL. PAL is being designed to arrange various kinds of events with concern for the who, what, when, where and why of that event. The goal for PAL is to permit a speaker to interact with it in English and to use extended discourse to state the speaker's requirements. The portion of the language system discussed in this report disambiguates references from discourse and interprets the purpose of sentences of the discourse. PAL uses the focus of discourse to direct its attention to a portion of the discourse and to the database to which the discourse refers. The focus makes it possible to disambiguate references with minimal search. Focus and a frames representation of the discourse make it possible to interpret discourse purposes. The focus and representation of the discourse are explained, and the computational components of PAL which implement reference disambiguation and discourse interpretation are presented in detail.

Keywords: reference disambiguation, discourse interpretation, discourse purposes, natural language, focus, frames.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under the Office of Naval Research under Contract Number N00014-75-C-0643.

CONTENTS

1. Introduction

Every discourse in English consists of one or more sentences which create a general context of people, places, objects, times and actions. The speaker of the discourse generally will not relate references from one sentence to the previous in any direct fashion nor indicate how the requests or assertions of each sentence in the discourse are connected. For the hearer to interpret the speaker's discourse and decide what the speaker is requesting or asserting, the hearer must complete two tasks, among others: (1) disambiguate the referential terms for their inter-sentential and extra-sentential links, and (2) determine the purpose of each sentence in the discourse. The first of these two tasks makes it possible to know what entities the speaker is referring to. The second task results in establishing a connected discourse and understanding what the speaker wants to communicate. Interpreting the discourse purposes of various sentences explains why D1 is acceptable below (even though D1-2 does not mention the party) while D2 is unacceptable. A theory of reference disambiguation will explain the disambiguation of *his* to Bruce and not to Mike, in D3.

D1-1 John is having a party at his house.
2 I think the guest of honor is Mary as they are going to announce the publication of Mary's book.

D2-1 Henry wants to meet with Harold.
2 Sing a song before 3 on Thursday.

D3-1 I want to have a meeting this week.
2 Bruce$_i$ will be the guest lecturer.
3 He will speak on slavery in ant colonies.
4 Mike wants to read his$_i$ report before the talk.

An explanation of these phenomena underlies the research being conducted at the MIT AI lab on PAL. While PAL is designed to understand the English form of requests for arranging various events, the design depends upon a theory about how to interpret a speaker's[1] extended discourse. PAL acts as a model of a hearer in these discourse situations. Two problems that must be solved before PAL can understand requests in extended discourse are referential disambiguation and discourse purpose interpretation. This paper reports on progress on these two problems.

A sample scenario of what PAL is designed to do is given in D4 below.

D4-1 I want to schedule a meeting with Dave.
2 It should be at 3 p.m. on Thursday.
3 We can meet in his office.
4 Invite Bruce.

---

1. I will use the term *speaker* to refer to the producer of a spoken or written discourse and *hearer* to refer to the receiver of the discourse.

To understand this discourse, PAL must have several natural language skills:

a. parsing for the syntactic structure.
b. interpretation of predicate-argument relations.
c. mapping of the words of each sentence to a representation used by the underlying database and programs.
d. disambiguation of the referential terms.
e. interpretation of each sentence for its discourse purpose.

The first two of these skills constitute the parser and case frame interpreter developed by Mitch Marcus. The representation mapping was developed by the author. These three modules are discussed in Marcus [1978]. To present a clearer picture of what PAL must be able to do, consider a sentence by sentence interpretation of the above dialogue.

I want to schedule a meeting with Dave.
PAL interprets an internal representation of the speaker as referent of "I," and an internal representation of "David McDonald" as the referent of "Dave."
PAL creates a new internal representation with features to be discussed later to be the referent of "a meeting."
PAL interprets "want to schedule a meeting" to be a request for a scheduling operation which may extend over several sentences.
PAL interprets the whole sentence to be asserting that the meeting has two participants, the speaker and Dave McDonald.

It should be at 3 p.m. on Thursday.
PAL interprets "it" as co-referring to the meeting under discussion.
PAL disambiguates the time phrase to a frame form used by the scheduler.
PAL interprets the sentence as asserting additional information about the meeting at hand.

We can meet in his office.
PAL determines that the speaker and other participant are the co-referent of "we."
PAL finds in its internal representations of things, an entity which "his office" can refer to.
PAL accepts the sentence as providing more information about the meeting at hand and asserts that fact.

Invite Bruce.
PAL finds an internal representation of the person referred to as "Bruce."
PAL determines that the ellided event which Bruce is to attend is the meeting under discussion.
PAL accepts the invite command as asserting another participant of the meeting.

<end of discourse>
PAL interprets the scheduling reuqest as complete and carries out the scheduling command with the meeting as it has been specified in the discourse.

In order to perform these tasks, a theory about the nature of discourse and some of its components has been developed and will be reported on here. Following that discussion, a closer look at the rules used by an implemented running version of PAL will be discussed.

2. Definition of Discourse

First, a "discourse" must be defined. I take a discourse to be any connected piece of text or spoken language of more than one sentence or independent sentence fragment. Ideally, every discourse is about some central concept which is then elaborated by the clauses of a discourse. Speakers often produce discourses which fail to meet this specification because they talk 1a) about several concepts without relating them or 1b) without informing the hearer that several concepts will be discussed at once or 2) because there is no central concept in their discourses. However, this idealization will serve to introduce some important terms. Multi-concept discourses do occur, and can be described using an approach which is a generalized version of that presented in this paper. Some cases of multi-concept discourse are discussed in Bullwinkle [1977]. However, the theory presented here has been tested in a running implementation of PAL, and this paper is restricted to that tested model.

In previous work [Winograd, 1971; Rieger, 1973; Charniak, 1972] various structures for referencing were assumed. Winograd used lists of entities of the same semantic type and chose referents for anaphoric terms based on recency and likelihood in the proper semantic class. His mechanism was too simple and failed to account for numerous anaphoric cases as well as being limited to objects in a closed world. Rieger postulated memory structures from a conceptual dependency representation of the sentences of a discourse. The memory structures were used to infer other information that could be unified to determine co-reference. His algorithms suffer from the explosive number of inferences that can be made from each memory structure. Charniak supposed that there were large collections of inference rules, called demons, which knew what to do with a small piece of the total knowledge, and which fired whenever that knowledge was encountered. This theory represents overkill; if one could have as many demons as Charniak supposed and get them to fire when that knowledge occurred, the mechanism could be used to predict co-referentiality of referencial terms. However, controlling the multitude of demons is difficult[2], and furthermore one cannot imagine how such a collection of knowledge is learned in the first place.

To interpret definite noun phrases and anaphors, a different approach is taken in PAL. It is assumed that discourse contains a structure, which when represented, can constrain the interpretation of referential terms. From the discourse structure, rules have been discovered which govern the acceptability of referential terms in different discourse situations. The interpretation of references is not strictly deterministic; it is like knowing

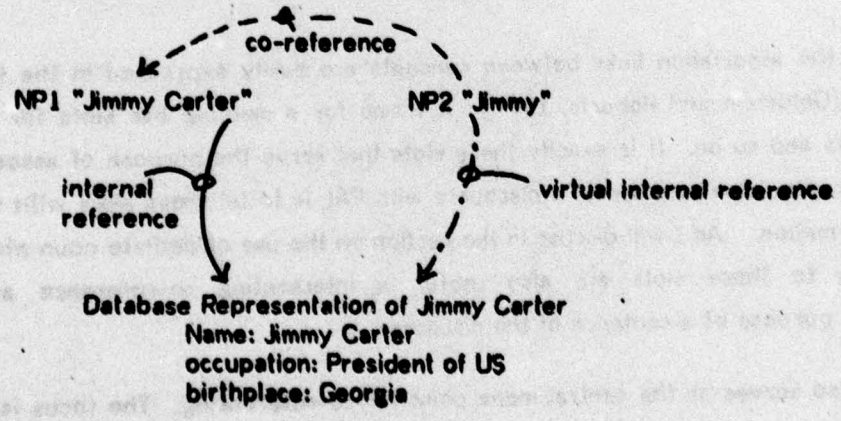
---

2. Rosenberg [personal communcation] has created a device called sentinels which may partially solve this problem.

which of several places to look in the discourse for a co-referent and trying out the term found there.

The theory underlying PAL distinguishes two kinds of referring. The first is an internal reference between a noun phrase and some pre-existing database object. That database object represnts a real world entity. In Figure 1 below internal reference links the noun phrase NP1 "Jimmy Carter" to a representation of Jimmy Carter (who is described as president of the US, etc.). How that database object refers to the real world is the classical semantic problem of reference (cf. Kripke [1972] among others) and is beyond the scope of this work. The other kind of referring is co-reference. Co-reference links a noun phrase to another noun phrase. The two noun phrases are said to co-refer, and both refer to the same database object. In Figure 1, the dashed link from NP2 "Jimmy" to NP1 is a co-reference link. The dot-dash link from NP2 to the database object is a virtual internal reference link which results from the co-reference link from NP2 to NP1 and from the internal reference link from NP1 to the database object. Internal reference and co-reference links are distinguished because co-reference links can be established more easily using discourse structure. In the remainder of this paper when I speak of internal reference, I will drop the phrase "internal" and use only "reference."

Fig. 1. Reference Links Between Noun Phrases



3. The Concept of Focus

The central concept of a discourse may be elaborated by several sentences of the discourse and then either discontinued in favor of a related concept, or dropped in favor of a new concept. This central concept of a discourse is called the discourse focus or simply the focus. This term was first used by Grosz [1977]. A simple example of focus is *meeting* in D4 repeated below:

D4-1 I want to schedule a meeting with Dave.
2 It should be at 3 p.m. on Thursday.
3 We can meet in his office.
4 Invite Bruce.

All four sentences give information about the focussed entity. The focus is what makes a text or a set of utterances a discourse.

In this work the focus is assumed to be a concept to which other concepts are associated. Some of the association links are "built-in" in the sense that they exist previous to the discourse. For example with *meeting*, built-in association links include that a meeting has a time, a place, a set of participants, and a topic of discussion. These association links are distinguished in the sense that the concept has explicit links to these concepts while no explicit links exist to other concepts such as color, cost or age. The discourse often serves the purpose of specifying more about the concepts linked to a focus. In D4-1, there is certain information about who the participants are, while D4-2 specifies the time. D4-3 causes the hearer to infer that the office is a place for a meeting, because the focus *meeting* has a place associated with it, and because PAL expects to be informed about the concepts associated to a meeting.

In PAL the association links between concepts are easily expressed in the frames structure of FRL [Goldstein and Roberts, 1977]. A frame for a meeting has slots for times, places, participants and so on. It is exactly these slots that serve the purpose of association links to other concepts. One purpose of a discourse with PAL is to fill those slots with values and required information. As I will discuss in the section on the use of definite noun phrases, the values given to those slots are also useful in interpreting co-reference and in understanding the purpose of a sentence of the discourse.

Focus also serves as the central index point for co-referencing. The focus is what is going to be talked about in the discourse. When it is introduced, it is new information. Thereafter it is the given information, and more new information is added to it. Knowing what the focus is helps determine co-reference relations because old information can be pronominalized while new information cannot. If a focus is seen not just as an entity by itself

but connected to other entities, focus indicates how those entities can be co-referents as well. In D4-(2-4), the focus of *meeting* can be used to determine the co-reference of *it*, *we* and *his* of *his office*: *it* must co-refer to the focus, *we* to those individuals associated to the focus who include the speaker, and *his* to an individual associated to the focus who is not the speaker and has male gender. The focus is used as an access function for retrieving the co-referent of a particular noun phrase. Later in this paper, rules governing the use of anaphora by means of the focus of the discourse will be discussed.

In the current version of PAL, focus is chosen as the first noun phrase following the verb if one exists, else the subject is used as focus. This method of choosing focus is adequate for current PAL discourses but not sufficient for the most general case. See Sidner [forthcoming] for a full discussion of focus choice. Once a focus is chosen, it can be used in succeeding sentences to determine the co-reference of pronouns or definite noun phrases as well as to check to see if the discourse is still connected. A sentence like (1a) below followed by (1b) is a disconnected discourse because the co-referential terms in (1b) are unrelated to the focus of (1a) based on the association links present in the database.

(1a) I want to meet with Henry.
(1b) Give me an ice cream cone.

The focus of the discourse can be changed while maintaining a connected discourse. The chief means are end of discourse remarks and focus-shift. End of discourse remarks can be explicitly stated ones like "That's all," or implicit ones, such as the act of simply ending the input stream. A less reliable, implicit marking of the end of discourse is to use a sentence with unrelated co-referential terms. In the case above, (1a) followed by (1b) could be assumed to be two separate discourses. This case is less reliable because it is impossible to tell if the speaker assumes that the ice cream cone is related (as is often the case with a non-ideal speaker) or whether the speaker intends to change the discourse to a new one. At present PAL does not accept this kind of abrupt discourse change; instead PAL indicates that such a sentence is not intelligible in the discourse. A more sophisticated PAL might request that the speaker explain how it is that (1b) is related to the discourse.

The other means of changing the focus I call focus-shift. A discourse may expand various aspects of a focus and then choose one aspect of the focus to describe in detail. For example, in a discourse about meetings, we may want to spend several sentences specifying the time for the meeting, why that time is best and so on. When time is being discussed, one would like to know that the focus has changed so that assertions or requests can be taken to be about time. However, the meeting focus may be brought back into the discussion later. To

maintain both foci, the meeting focus is stacked for later use.[3] Detecting this focus change is the process of focus-shift.

Focus shifts cannot be predicted; they are detectable only after they occur. To detect the focus shift, the focus shift mechanism takes note of new phrases in sentences following the introductory sentence. Any new phrase is a potential focus. an anaphoric term in a sentence which follows the potential focus sentence may co-refer to either the focus or the potential focus. If the potential focus is an acceptable co-referent, it is the co-referent of the anaphoric term, and the focus shifts to the potential focus. The choice of office as co-referent of *it* in D5-3 results from focus-shift. The co-referent of *it* to meeting in D5-3' results from the rejection of the potential focus office as the co-referent.

D5-1: I want to schedule a meeting with George, Jim, Steve and Mike.
2 We can meet in my office.
3 It's kind of small, but the meeting won't last very long anyway.
3' It won't take more than 20 minutes.

Rejection of a co-referent results from semantic information about the type of verb and the type of semantic entities it accepts. Semantic information has been proposed for use with co-reference (see Winograd [1971], among others). PAL uses this information only to reject or confirm choices made by the focus and focus-shift mechanisms, rather than to suggest classes of co-referents.[4]

4. Modules of PAL

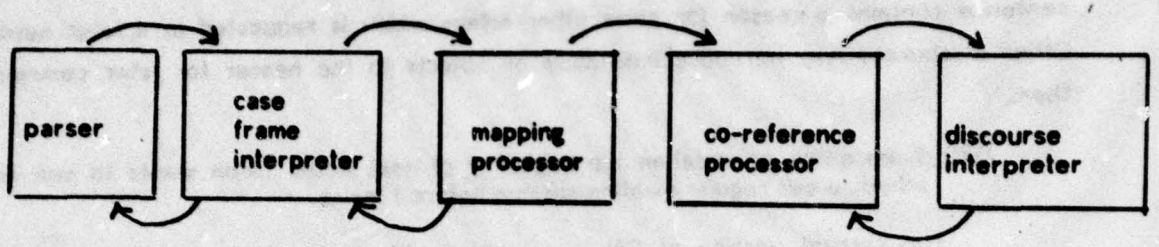
The preceding description of co-reference interpretation has been incorporated into a series of modules for PAL. These modules are depicted in Figure 2 below. The arrows represent flow of control between modules.

Each English sentence presented to PAL by a speaker is interpreted via a parser, case frame interpreter and representation mapping program [Bullwinkle, 1976; Marcus, 1978] into a set of FRL frames. The sentence "Schedule a meeting in my office," is represented by the following simplified frames (slot and slot values are listed also).

---

3. Grosz [Deutsch, 1975] gave the first specification of discourse shifts using the concept of focus. These are discussed further in Grosz [1977].
4. The mechanism of focus-shift is discussed in more detail in Bullwinkle [1977], where the term "sub-topic shift" is used.

Fig. 2. Modules of PAL



---

frame	=	schedule201
a-kind-of	=	schedule
type	=	"imperative"
actor	=	PAL
event	=	meeting203
frame	=	meeting203
a-kind-of	=	meeting
place	=	office207
determiner	=	"a"
frame	=	office207
a-kind-of	=	office
determiner	=	my209
frame	=	my209
a-kind-of	=	my

Given these frames, PAL is expected to determine what my209, office207, and meeting203 co-refer to. PAL also must decide what the purpose of an imperative scheduling request (represented by schedule201) is relative to its database collection of actions. Each of these modules will now be discussed in detail.

5. Interpretation of Discourse Purposes

To interpret discourse purposes, a discourse module creates a model of the discourse and controls the process of focus identification. Since the beginning, middle and end of a discourse each require different actions by the PAL scheduler, the discourse component models each differently. The first sentence of the discourse is assumed to specify what the nature of the user's communication is. This is a simplified view of the real communication process. Many discourses do not simply state their object and then elaborate the relevant information. Instead many speakers begin a discourse as in D6 below in which the first

sentence contains a reason for some other action, which is requested in a later sentence. Other discourses may introduce individuals or objects to the hearer for later comment on them.

D6: I am going on vacation the beginning of next week. John wants to see me, so schedule our regular meeting session before I leave.

The current version of PAL uses the simplified view of discourse to choose a discourse purpose. Introductory sentences are assumed to be making some sort of request. The PAL discourse module chooses which request on the basis of the verb and any associated modals, or on the basis of verbs of desire (*want, wish, would like*) and the verb complement. A request consists not only of the request type, but of some object which the request is about (intransitive verbs are not relevant to PAL since telling PAL to laugh, run or groan is inappropriate). The focus of the discourse is used for this purpose. This choice is plausible not only because the focus is closely associated with the object of the verb, but also because a discourse centers discussion on some particular entity, and that entity is captured by the focus.

Once a focus has been designated, sentences occurring in mid-discourse are assumed to be about the focus until the co-reference module predicts a focus-shift and/or until the verbs used are inconsistent with the discourse request. Mid-discourse sentences often do not explicitly co-refer to the focus as has been shown previously in D1 and D4; they may contain an implicit focus co-reference. Use of focus for co-reference disambiguation has the added benefit that sentences containing implicit focus co-references are easily recognized by the discourse component. Once an implicit focus relation is established, the module can go onto predictions of focus shift. Knowledge that the speaker is co-referring to the focus, either explicitly or implicitly, makes possible the prediction that the discourse is not yet complete, and the prediction that the speaker is making a coherent request. Since neither prediction can be assumed trivially true, the focus is important to the communication process.

In addition to the focus, the discourse module contains knowledge allowing the module to decide if the verb of a new sentence is consistent with the discourse request. Thus in D7 below, the second sentence uses a verb that is consistent with the scheduling request while in D7', the verb is odd.

D7: Henry wants to meet with Harold. Choose a time before 3 on Thursday.
D7': Henry wants to meet with Harold. Sing a song before 3 on Thursday.

The knowledge needed to predict consistency is represented in the frames database in two ways. First the frame for the discourse request contains information about what other requests can be sub-requests of the discourse. Second a set of mapping frames contain information which determine how a verb can be interpreted as making a certain request. For

example, the verb *be* can be associated with scheduling and re-scheduling activities. However, the intention of the speaker in a sentence like (2) is different within the context of a scheduling or a re-scheduling request.

(2) The time should be 3 pm.

In a scheduling context, (2) can be interpreted to request that the time be established as 3 pm while (2) in re-scheduling can have an interpretation of changing the time from whatever it was to 3 pm. PAL captures the intention of the speaker relative to a request context by an inference mechanism which is a matcher that determines that (2) represented as a frame[5] can be associated with scheduling requests by a simple mapping between two frames. This correspondence coupled with the use of focus makes it possible to understand (2) as part of a discourse.

In addition, the mapping functions tell how to interpret the current sentence into one of the commands which the scheduler can perform. Included in this process are how to map the slots of one frame into a frame which is the scheduling action. For example, the verb frame for "We can meet in 823" is mapped from a "meet" frame into a frame called "assert" with a slot for the object asserted, which is the focus, and a slot for what is asserted about that object, in this case the place as 823.

The end of a discourse is currently interpreted as being the end of the speaker's input stream. A more sophisticated means of interpreting discourse end is possible, though not implemented, given the focus mechanism: when the needed slots of the focus are filled, the speaker can be considered to have finished this discourse. Upon sensing the end of the discourse, the discourse module informs the scheduler that it can carry out the action requested at the discourse beginning. At first glance this may appear as if the discourse request specified at the beginning is ignored in favor of other requests. In fact the initial request is used in interpreting mid-discourse sentences. However, many discourse actions like scheduling require that the action of scheduling be delayed until all the necessary information for scheduling is presented. This process normally cannot be stated in a single sentence, and a whole discourse is needed to fill in the request. In this fashion the discourse module reflects the fact that a discourse consists of many sub-discourses centered around individual entities and which are opened and closed by focus shifting or finishing discussion of the current focus.

PAL is similar to the GUS system [Bobrow et al, 1977] because it expects a discourse to provide information about the slots of a frame. GUS permits user initiative

---

5. A frame is not taken as the meaning, in the classical semantic sense, for (2); PAL makes no claims about this sense of meaning.

although it is unclear what the extent of this initiative is. GUS does not seem to allow for user initiative of the discourse requests. Since PAL expects full user control over all parts of the discourse, PAL needs a complete description of the discourse and its focus. PAL's use of focus also presents a complete theory of the kinds of co-reference problems raised by the GUS system.

6. Co-reference Disambiguation

There are two sub-modules for co-reference interpretation in PAL, the sentential and inter-sentential co-reference modules. The inter-sentential co-reference sub-module chooses co-references for referential terms in the discourse once the focus is identifed. The task of determining co-reference varies depending upon the presence or absence of previous discourse. When there is previous discourse, co-reference interpretation depends largely on the focus. For simple[6] definite noun phrases, PAL assumes either the focus is the direct co-referent of the definite noun phrase or the focus contains a slot that is the co-reference of the definite noun phrase. This assumption needs modification since some definite noun phrases are used to refer outside the context of the discourse. For example, when trying to schedule a meeting, if the speaker says (3), the definite noun phrase co-refers to an entity associated with the meeting under discussion; that association is reflected in the frame slot structure of FRL.

(3) The best place is my office.

However, if the speaker says (4), *the conference room*, i.e. that particular conference room which the speaker has in mind, is not associated with meetings in general, and so the focus does not point out the co-reference.

(4) We ought to meet in the conference room.

However, by searching the focus, the lack of a connection can be noticed, and a reference from the database can then be considered. In this way, the focus acts as an access function, but only for those co-referential terms related to the previous sentences of the discourse.

PAL uses database search with growing contexts of reference to choose reference for other kinds of noun phrases which refer to entities outside the discourse. Growing a context is accomplished using the immediate set of frames from the first sentence and recursively creating larger sets from the slot values of those frames until the frame with the name in question is found. The context growing mechanism reduces search from a more global search strategy, and helps control potential ambiguities that exist due to multiple possible

---

6. A simple definite noun phrase is a definite noun phrase containing no relative clauses. At present PAL interprets only such noun phrases.

references in the database. This same method could be used for definite noun phrases that refer outside the discourse.

Use of the focus is actually somewhat more complex since the definite noun phrase may be a co-reference to the potential focus of the discourse. Should a definite noun phrase co-refer to the potential focus, the discourse module pushes the current focus to a focus stack and takes the potential focus as the new focus. The pushed focus is available for later use in the discourse. The current inter-sentential sub-module does not interpret definite noun phrases used generically. The focus can be used for these cases as well (see Sidner, [forthcoming]), but the details of this process are not included in the current version of PAL.

The inter-sentential co-reference sub-module also determines the co-reference of personal pronouns. For the pronouns of first person plural (*we, us*), two choices can be made. First the sub-module can choose the focus as the direct co-referent of the anaphor. Second the sub-module can choose a set of co-references from a particular slot of the focus. That slot must contain co-references including the speaker of the discourse. For *he/she*, and its object forms, the focus is chosen as a direct co-reference. Using the focus as co-referent explains the anaphoric co-reference in D8 of *his* to Bruce and rather than Mike. When the focus is not the co-referent, a co-referent stipulated by the co-reference rules of the sentential co-reference sub-module, discussed below is used. Finally if neither is acceptable, entities associated with the focus are checked for co-reference. This sub-module predicts misuse of *he/she* pronouns if no co-references are found from this process or if more than one results from the last step in the process.

The interpretation of co-reference for *he/she* pronouns needs to be expanded to include consideration of potential focus since in D8 below, *his* co-refers to Bruce and not to Mike.

D8: I want to have a meeting this week. Bruce$_i$ will be the guest lecturer. Mike wants to read his$_i$ report first.

It appears that the focus and potential focus ought to be checked for co-reference to such pronouns before sentential co-reference rules are used. However, further experimentation with such cases is needed to confirm this aspect of co-reference.

For the co-reference of *it*, the inter-sentential co-reference sub-module chooses a co-referent either from the focus, the potential focus or from predictions from sentential co-reference rules, which are discussed below. This choice strategy is not entirely adequate because recency appears to play a role in the co-reference choices for *it*. Recency rules are discussed in Sidner [forthcoming], and could be included in a future version of PAL. The

inter-sentential co-reference sub-module uses the semantic constraints placed on the pronoun by the verb in a few instances; this portion of PAL could be expanded greatly. Co-reference rules for *they* work similarly to those for *it* with consideration that the speaker cannot be included in the co-reference set.

When no previous discourse exists, PAL's sentential co-reference sub-module uses the co-reference rules of Lasnik [1976] to choose co-references. The rule is stated as follows: If a noun phrase, NP_1, precedes another noun phrase, NP_2, and NP_2 is not a pronoun, and further if the minimal cyclic node dominating NP_1 also dominates NP_2, then NP_2 and NP_2 *are disjoint in reference. The expression "disjoint in reference" is taken to mean have no* references in common, thereby blocking the co-reference of *Bob* and *Tom* to *they* in (5).

(5) They assume that Bob will talk to Tom.

By using Lasnik's rule, disjoint references of a noun phrase in a sentence can be chosen, as well as a list of acceptable co-references for the noun phrase. This information is recorded in the frame presenting the noun phrase. As pointed out by Reinhart [1976], Lasnik's rule fails to predict the disjoint references in sentences like (6) and (7) below, but these cases are not problematic given inter-sentential co-reference rules because other rules will predict the co-reference for the pronouns first.

(6) Near Dan, he saw a snake.
(7) For Ben's wife, he would give his life.

In addition to the use of a co-reference rule, the sentential sub-module determines the referents of proper names. Using the collection of frames which make up the discourse, a frame containing the correct first (and if given, last) name can be found. Should the immediate discourse fail to produce the name referent, a larger context can be grown from the slot values and from the slot defaults of the frame representing the focus. The same context growing mechanism used for definite noun phrases is used. By this process of context growing, ambiguous uses of names like *John* can be avoided. *John* will refer to that person most closely associated with the discourse. If more than one frame for the name John is found, the context growing process predicts that the speaker has used the name ambiguously. Context growing has been effective in a limited number of cases tested so far, although a database with more potential ambiguities would further test this sub-module.

7. Extensions

The current PAL can be expanded in many directions. Some of the necessary developments of its co-reference capabilities have already been discussed. Significantly, these capabilities do not require extensive new theoretical appparatus; the focus of discourse and structure of FRL can sustain the needed improvements. In discourse interpretation PAL must be extended to interpret discourses which define new people, places, events, actions and like objects as well as to interpret preferences of users and purposes for various activities. These extensions not only will make PAL a more useful system, but also they encompass a set of tasks useful for other interactive programming domains. Experimentation on the discourse module of PAL is need to incorporate these new capabilities.

8. Acknowledgements

The author wishes to thank Gretchen Brown, Ira Goldstein and Steve Rosenberg for their comments and suggestions on drafts of this paper.

9. References

Bobrow, D., R. Kaplan, M. Kay, D. Norman, H. Thompson, T. Winograd, [1977] *GUS, A Frame-Driven Dialogue System*, Artificial Intelligence, Volume 8, Number 2, April.

Bullwinkle, C. [1976] *The Semantic Component of PAL: the Personal Assistant Language Understanding Program*, MIT AI Laboratory Working Paper, March.

Bullwinkle, C. [1977] *Levels of Compexity in Discourse for Anaphora Disambiguation and Speech Act Interpretation*, Proceedings of the Fifth International Joint Conference in Artificial Intelligence, August 1977; also M.I.T. A.I. Lab Memo 413.

Charniak, E. [1972] *Toward a Model Of Children's Story Comprehension*, M.I.T. A.I. Lab TR-266.

Deutsch, B. [1975] *Establishing Context In Task-Oriented Dialogues*, Proceedings of the 13 Annual Meeting of ACL, AJCL Microfiche 35.

Goldstein, I. P. and R. B. Roberts. [1977] *NUDGE, A Knowledge-based Scheduling Program*, M.I.T. A.I. Lab memo 405.

Grosz, Barbara [1977] *The Representation and Use of Focus in Dialogue Understanding*. Stanford Research Institute Technical Note 151, Menlo Park, California

Kripke, Saul A. [1972] *Naming and Necessity*. in Semantics of Natural Language, Davidson and Harman (eds.) Reidel Publishing Co., Boston.

Lasnik, Howard. [1976] *Remarks on Co-reference*, Linguistic Analysis, Volume 2, Number 1.

Marcus, Mitchell [1978] *Progress Report on the Parser and Semantics of PAL*, M.I.T. A.I. Lab memo forthcoming.

Reinhart, Tanya [1976] *The Syntactic Domain of Anaphora*, unpublished Ph.D. dissertation, Department of Foreign Literature and Linguistics, M.I.T.

Rieger, Charles J. [1974] *Conceptual Memory: A Theory and Computer Program for Processing the Meaning Content of Natural Language Utterances.* Stanford Artificial Intelligence Lab Memo AIM-233.

Sidner, C. [forthcoming] *A Computational Model of Co-reference Comprehension in English.* Ph.D. dissertation, M.I.T.

Winograd, Terry [1971] *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language.* M.I.T. dissertation.